



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Journal Pre-proof

Genomic evolution of SARS-CoV-2 in Reunion Island

David A. Wilkinson, Alizé Mercier, Magali Turpin, Marie-Alice Simbi, Jonathan Turpin, Camille Lebarbenchon, Maya Cesari, Marie-Christine Jaffar-Bandjee, Laurence Josset, Luce Yemadje-Menudier, Bruno Lina, Patrick Mavingui



PII: S1567-1348(22)00178-2

DOI: <https://doi.org/10.1016/j.meegid.2022.105381>

Reference: MEEGID 105381

To appear in: *Infection, Genetics and Evolution*

Received date: 26 July 2022

Revised date: 18 October 2022

Accepted date: 25 October 2022

Please cite this article as: D.A. Wilkinson, A. Mercier, M. Turpin, et al., Genomic evolution of SARS-CoV-2 in Reunion Island, *Infection, Genetics and Evolution* (2022), <https://doi.org/10.1016/j.meegid.2022.105381>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2022 Published by Elsevier B.V.

Genomic evolution of SARS-CoV-2 in Reunion Island

David A Wilkinson^{a,c,*}, Alizé Mercier^b, Magali Turpin^a, Marie-Alice Simbi^{a,c}, Jonathan Turpin^{a,c}, Camille Lebarbenchon^a, Maya Cesari^c, Marie-Christine Jaffar-Bandjee^d, Laurence Josset^e, Luce Yemadje-Menudier^b, Bruno Lina^e, Patrick Mavingui^{a,*}

^aUniversité de La Réunion, UMR Processus infectieux en milieu insulaire tropical (PIMIT) CNRS 9192, INSERM 1187, IRD 249, Sainte-Clotilde, La Réunion, France

^bSanté publique France, Saint-Denis, La Réunion, France

^cGIP CYROI, Saint-Denis, La Réunion, France

^dLaboratoire de Microbiologie, Centre Hospitalier Universitaire de la Réunion, Saint-Denis, France

^eCNR des Virus des Infections Respiratoires, Institut des Agents Infectieux, Hospices Civils de Lyon, Lyon, France; Virpath, Centre International de Recherche en Infectiologie, Université de Lyon, Inserm U1111, CNRS UMR5308, École Normale Supérieure de Lyon, UCBL, Lyon, France.

Corresponding authors: Dr David A Wilkinson, Dr Patrick Mavingui

*email: david.wilkinson@univ-reunion.fr; patrick.mavingui@cnrs.fr

Abstract

Island communities are interesting study sites for microbial evolution during epidemics, as their insular nature reduces the complexity of the population's connectivity. This was particularly true on Reunion Island during the first half of 2021, when international travel was restricted in order to mitigate the risk for SARS-CoV-2 introductions. Concurrently, the SARS-CoV-2 Beta variant became dominant and started to circulate at high levels for several months before being completely replaced by the Delta variant as of October 2021. Here, we explore some of the particularities of SARS-CoV-2 genomic evolution within the insular context of Reunion Island. We show that island isolation allowed the amplification and expansion of unique genetic lineages that remained uncommon across the globe. Islands are therefore potential hotspots for the emergence of new genetic variants, meaning that they will play a key role in the continued evolution and propagation of COVID-19 as the pandemic persists.

Keywords

SARS-CoV-2, evolution, genomics, Reunion Island

1. Introduction

Oceanic islands have historically been attractive environments for the study of the ecology and evolution of many organisms [1]. Recognized as “natural laboratories”, these peculiar ecosystems have stimulated research on colonization, diversification and specialization processes of a wide range of animal and plant species, and recently of human diseases [2]. Indeed, tropical islands are in the front line for disease emergence related problems; the islands of the Western Indian ocean, Caribbean and Pacific often shoulder this burden, as demonstrated by recent outbreaks associated to vector-borne diseases such as Chikungunya, Dengue and Zika [3–5].

Identifying the source, as well as understanding the mechanisms underlying transmission dynamics, and predicting the evolutionary trajectory of emerging viruses remains very challenging. Because of their peculiar features, oceanic islands provide relevant conditions for the study of such aspects. For instance, spatial isolation may represent a major barrier to the introduction of infectious agents. The discrete geographical nature of islands allows the identification and quantification of infectious agent migration processes associated with human travel and animal trade; and their small sizes also offer possibilities for rapid and efficient monitoring of epidemics. Oceanic islands could be considered as simplified ecosystems as compared to continental habitats, in which the quantification of the factors involved in diseases transmission may be easier to achieve, and where intervention measures may be more effective on smaller scales.

The COVID-19 pandemic, caused by SARS-CoV-2, has had an unprecedented impact on the global community since its suspected emergence from wildlife populations in China in 2019 [6,7]. According to WHO data to date (July 2022), more than 500,000,000 cases of COVID-19 have been declared across the globe, with more than 6,000,000 associated deaths. Over more than two years, patterns of human movement associated with both trade and travel have shaped the progression of the pandemic, with countries introducing movement restrictions at differing scales to try to mitigate the spread of the disease. However, all global communities have now been affected by COVID-19. This includes many small island communities, which have faced unique challenges dealing with the burden of the COVID-19 pandemic in insular and often geographically isolated contexts, sometimes with limited access to healthcare and medical resources.

In this study, we investigated SARS-CoV-2 genomic evolution in the island of Reunion, a French overseas territory located in the Western Indian Ocean. We hypothesized that population factors characterizing small islands could favour the selection and emergence of particular variants. But contrastingly, this effect may be diluted by continuous population mixing associated with numerous incoming and newly infected travellers. Based on 7,423 consensus genome sequences generated between the 1st of January 2021 and 15th April 2022, we investigate key evolutionary features of the epidemic in Reunion, paying particular attention to aspects of the epidemic that were unique (or otherwise uncommon) to this particular island setting.

2. Materials and Methods

2.1. Sample and data collection

COVID-19 positive nasopharyngeal swabs identified as part of routine diagnostic screening across Reunion Island were provided for genomic analysis. Epidemiological data relating to COVID-19 cases between 1 January 2021 and 15 April 2022 were collected through the national and regional surveillance system (SI-DEP).

2.2. Genomic data

Genomic data were generated using the ARTIC amplicon-tiling protocol for the Oxford Nanopore Technologies (ONT) sequencing platform (MinION) [8]. Briefly, RNA was extracted from nasopharyngeal swabs using the Qiagen RNA Mini extraction kit on the Qiacube pipetting robot. cDNA was generated using the Lunascript random hexamer reverse transcriptase protocol from New England Biolabs. Overlapping PCR products spanning the entire length of the genome were amplified in two independent reactions. The ARTIC v3 standard primer sets until October 2021, and the ARTIC v4 standard primer sets after this [9]. Changing primer sets corrected for mutations in the primer regions that were present in the Delta variant. PCR products were pooled and quantified using the DNA HS assay kit on a Qubit 4.0 (ThermoFisher). Native barcodes were attached to PCR products via ligation using the TA blunt DNA ligase before samples were pooled and purified using AmpPure XT magnetic beads (Beckman). The nanopore sequencing adapter

molecule was attached to the purified barcoded DNA using the Quick T4 DNA ligase kit from New England Biolabs before being purified for a second time using AmpPure XT magnetic beads. Sequencing of DNA libraries was performed on the MinION Mk1c device using R9.4.1 flowcells from Oxford Nanopore. Sequencing times varied depending on the number of samples included in each library, but overall, sufficient data were generated to provide a median read depth of 1,484x. Basecalling and demultiplexing were performed in high accuracy mode and specifying the "--require-both-ends" parameter using the most recent version of the Guppy software from Oxford Nanopore at the time of sequencing. Genomic assemblies were performed using the latest version of the Medaka assembly software at the time of sequencing. Genomes with more than 5 % missing data (1495 of 29903 bases), or a median read depth lower than 300x were excluded from genetic analyses.

2.3. Lineage assignment and phylogenetic analyses

Lineage assignments were performed using PAN-GOLIN (Phylogenetic Assignment of Named Global Outbreak Lineages) v3.1.2 [10], and the NextClade CLI v1.7.0 [11]. The output from NextClade was used for sequence alignment and to identify the presence of specific substitutions, deletions, and insertions for each genome.

Data manipulation and presentation were performed using the R software [12], with the packages "stats", "dplyr" [13], "ape" [14], "phytools" [15], "ggtree" [16] and "DECIPHER" [17]. Phylogenetic analyses were performed using the COVID-19 release of the IQtree software (version 2.1.3) [18]. Default settings to identify the best candidate substitution model, GTR+F+R3 according to BIC. Node support was calculated from 1000 bootstrap replicates. Treetime v0.7.0 [19] was used to estimate node divergence dates from the consensus maximum-likelihood tree output of IQtree. The outputs of treetime were used for all root-to-tip analyses. The tree was rooted relative to the original Wuhan isolate genome (NCBI accession number MN908947.3). International genomic data was extracted from the GISAID database [20] and all contributing laboratories are referenced via GISAID's EPI-SET tool (Supplementary Data 1).

3. Results and Discussion

3.1. Epidemiological context

The epidemiological aspects of the COVID-19 in Reunion Island have been partially addressed elsewhere [21]. Briefly, the number of observed cases of COVID-19 rose in Reunion Island at the beginning of 2021, following the Christmas-period holidays of 2020. This rise in cases was associated with the introduction of the Beta variant, which became the main circulating variant on the island by February 2021 (week 7) of 2021. Case incidence then remained stable for approximately 20 weeks at around 100 cases per 100,000 inhabitants until the introduction of the Delta variant, which resulted in a sharp rise in incidence prompting the rapid implementation of stricter containment measures. Curfew and social-distancing measures resulted in a steady decrease in case incidence, which was accompanied by an increase in the proportion of Delta-variant associated cases. By the end of September 2021 (week 39), the Delta variant had completely displaced the Beta variant as the main circulating lineage on the island (Figure 1). The first imported case of Omicron was observed in November (week 46), and autochthonous circulation of Omicron likely started around week 48 of 2021. The circulation of Omicron was accompanied by a sharp rise in the number of reported cases, which peaked at approximately 3,500 cases per 100,000 individuals in week three of 2022. By the end of the study period (week 15 of 2022) Omicron lineage BA.2 had entirely replaced Omicron lineage BA.1 and was essentially the only variant circulating on the island.

3.2. Lineage replacement and sequential waves of variant circulation

A total of 7,423 genomes with greater than 95% coverage at a median read depth of 1,484x (IQR 962 - 2003) were included in this study. Genomes originated from a pseudo-random selection of individuals that had undergone testing for COVID-19 infection in private laboratories across Reunion Island between the 1st of January 2021 and 15th April 2022. Weekly sequencing activity

generated genomic data for a median of 5 % (IQR 3.1 % - 10.7 %) of all declared COVID cases in Reunion.

Based on pangolin classification data, the major trend in lineage evolution was one of complete, or near-complete selective sweeps that lead to lineage replacement on at least four separate occasions (Figure 1). The first of these transitions has been documented elsewhere [21]; Briefly, non-variant-of-concern (VOC) lineages were replaced by Beta subvariants B.1.351 and B.1.351.2 around week 7 of 2021. Prior to this time, high proportions of imported cases were reported (Figure 1) that were associated with circulating European lineages. For example, subvariants of lineage B.1.177 were common, this lineage is thought to have originated in Spain and circulated predominantly around Europe towards the end of 2020 [22].

Among the non-VOC lineages identified in weeks 1-15 of 2021, we identified a strain that showed significant dissimilarity to all available sequences in the GISAID database. As a result, lineage B.1.622 was assigned to the pangolin database based on data from 71 observed cases. We detected B.1.622 in Reunion between weeks 03 and 15 of 2021, and its relative detected prevalence was highest in week 06, when B.1.622 sequences represented 13 % of all sequenced genomes. B.1.622 isolates were highly divergent from the original sequence obtained from Wuhan in 2019 (Figure 3), carrying an average of 24 synonymous and non-synonymous mutations relative to the reference (Table 1). This level of divergence is similar, for example, to that observed for VOC Beta variants at the same time. The GISAID online tool CoVizu^e was used to identify similar lineages from the global isolate database. Lineage B.1.622 was predicted to share a common ancestry with two other Pangolin lineages; B.1.474 which showed a widespread distribution across Europe, and B.1.629, a relatively uncommon lineage which was identified both in continental Africa and across Europe (Figure 3). All three lineages exhibited significant genetic dissimilarities to other known strains and each other, appearing at the ends of long branches from the B.1 parental root position of the global time-scaled tree. Interestingly, a single sequence of B.1.622 was also isolated in France, falling at a basal position close to the root node of the B.1.622 clade in phylogenetic analyses (Figure 3).

B.1.622 carried mutations characteristic of other major lineages from 2020 such as the D614G spike mutation that increases SARS-CoV-2 fitness [23]. In addition, B.1.622 was characterised by a specific genetic profile combining: five mutations in the polymerase (ORF1a:S428N, T2196N, M3621I and L3977V; ORF1b: S2312L); two mutations in the Nucleocapsid gene, Q70R, which is located in the SARS-CoV-2 specific T cell epitope [24], and S187L in the linker domain; one mutation in ORF9b, K67E. Moreover, there was a high density of mutations in the ORF3a gene including Q57H which was associated with a wave of infection in China and has been shown not to have a significant effect on SARS-CoV-2 fitness [25], as well as T14I and S216P which have both been described to alter the structure and function of the ORF3a-encoded protein [26]. The full list of amino acid substitutions is shown in Table 2.

Beta subvariants B.1.351 and B.1.351.2 co-circulated in Penang in relatively stable proportions for nearly 35 weeks before being displaced by Delta in the second major strain transition that occurred around week 29 of 2021. Interestingly, however, the majority of Delta strains in circulation could be attributed to one of two pangolin lineages, AY.40 (974 of 3098 isolates) and AY.43 (1092 of 3098 isolates). Arrival of AY.40 around week 27 resulted in a large peak in the number of infections at week 31 (weekly incidence of infection of 350 per 100,000). AY.40 infections dropped to much lower levels by week 41, and AY.40 was itself replaced by AY.43. Case numbers of predominantly AY.43 infections then increased steadily from weeks 41 to 51 (Figure 1).

The third major lineage replacement event occurred between weeks 52 of 2021 and 02 of 2022, when circulation of VOC Omicron BA.1 began during a period of time when AY.43 (Delta) infections were increasing. This likely had a synergistic effect on infection rates, resulting in a massive increase in case incidence to above 3,500 per 100,000 individuals in week 03 of 2022.

Omicron BA.1 remained the main detected circulating variant for only 8 weeks. From week 09 of 2022 Omicron BA.2 became the majority variant and had almost fully replaced Omicron BA.1 by week 12. It should be noted that due to the extremely elevated case numbers, only a small fraction of cases was sequenced during 2022, and that case travel history is absent for the majority of this period due to the concurrent saturation of the health system.

3.3. Different dynamics and genetic characteristics of VOCs

Overall dynamics were similar for VOCs Beta, Delta and Omicron, with each variant successfully replacing its predecessor rapidly after the establishment of autochthonous circulation on the island. However, differences in the relative selective advantages between competing lineages resulted in different durations of the periods of lineage replacement. Beta became the majority variant (>50% of sequenced genomes) after seven weeks of circulation, compared to five weeks for Omicron and two weeks for Delta (Figure 4).

Genetic diversity profiles of each VOC were highly contrasting. Beta lineages showed low diversity after importation, with overall strain diversity increasing almost linearly over the entire duration of its circulation in Reunion (Figure 4). In contrast, high strain heterogeneity of Delta variants was observed rapidly after its arrival, both in terms of the number of predefined pangolin lineages detected (Figure 1) and overall genetic diversity (Figure 4). Interestingly, the overall genetic diversity of Omicron variants was also observed to increase linearly over time but remained lower than that of both Beta and Delta variants (Figure 4) despite a large number of predefined Pangolin lineages being detected (Figure 1), reflecting a likely high rate of importation. This may be indicative of negative selection limiting Omicron divergence.

We examined the amino acid substitutions of all VOCs and lineages that had demonstrated significant selective advantage in Reunion between 2021 and 2022. All substitutions that were conserved within VOCs are shown in Table 2, whereas all substitutions that distinguish between locally important subvariants of VOCs are shown in Table 3. Interestingly, there were several observations of the same substitution being acquired independently by different lineages, potentially indicating positive selection for these mutations. Local lineage B.1.622 shared ORF3a:Q57H with Beta variants B.1.317 and B.1.317.2. Beta variants shared spike mutation S:N501Y all Omicron variants, and spike mutation S:K417N was also shared between Beta variants and Omicron BA.2 subvariants. Both of these mutations have been linked to immune escape mechanisms [27], and S:K417N was additionally observed in some Delta subvariants belonging to pangolin lineage AY.4.2. Interestingly, strain AY.43 (Delta) which replaced similar

variant AY.40, shared the ORF1a:T3255I with all Omicron variants as well as spike gene mutation S:G142D with Omicron subvariant BA.2. However, the S:G142D mutation has been identified as a potential sequencing artefact resulting from the use of the ARTIC V4 primer sets [28].

As an overseas department of France, Reunion Island shares a strong link with continental France, and is also a travel hub between Europe, Africa, and Asia. As such, one of the key questions surrounding the epidemiology of SARS-CoV-2 in Reunion Island is why the Alpha variant did not become dominant during 2021, when it was the majority variant in continental France and in Europe. We observed multiple importation events of Alpha variants to Reunion (Figures 1 and 3), but the proportion of Alpha cases was never observed to increase significantly (Figure 4). Studies in other island communities have reported importation bottlenecks that result in the delayed entry of otherwise common variants. This, for example, resulted in unique variant profiles in the Canary Islands despite strong ties with continental Spain [29] and the island of Ilhabela despite strong ties with continental Sao Paulo (Brazil) [30]. It may therefore be that Alpha was not imported sufficiently frequently for it to become a dominant variant or it may be that imported Alpha variants did not possess a sufficient selective advantage to outcompete others in a population where Beta variant circulation was already dominant.

It is not known if the Beta variant possesses selective advantages that may account for it becoming the dominant lineage on Reunion, although it has been suggested that Beta variants exhibit partial immune evasion meaning that they could have a selective advantage in populations where cumulative incidence exceeds 20% [31]. As the cumulative number of cases reported in Reunion at the beginning of 2021 represented only 1% of the total population, this is unlikely to explain the prevalence of Beta. The linear increase in genetic diversity of Beta variants over time and associated tree topology (Figure 3) suggest continued genetic drift amongst local transmission chains in a large naive population, resulting in local strain diversification. Simultaneously, low numbers of imported cases were observed during the main period of Beta circulation, due to movement restrictions and confinement measures (Figure 1).

Of particular interest, the most frequently detected Beta subvariant in Reunion Island, pangolin lineage B.1.351.2, has never been reported to circulate significantly outside of the Indian Ocean region, but was frequently reported in Reunion Island, Mayotte and Madagascar over the same period. Local population movements thus appear to have generated an insular community which allowed this otherwise rarely-observed lineage to circulate and diversify on a large scale.

The transmission advantage of the Delta variant is largely explained by its ability to replicate faster compared to many other VOCs [32]. However, the genetic diversity of Delta observed in Reunion Island (Figure 4) as well as relative increases in the proportion of imported cases surrounding the period of Delta circulation (Figure 1) suggests a very different dynamic when compared to Beta circulation. Numerous subvariants of Delta were likely imported very frequently, and a resulting two-phase selective sweep was observed with AY.40 and AY.43 becoming dominant variants. The observed dynamic is, again, different for Omicron; which was imported on multiple occasions, has a selective advantage explained by increased replication rates, immune evasion and increased reinfection rates [32], circulated in high levels but still retained a low level of genetic diversity compared to other variants.

3.4. Substitution rate heterogeneity

We calculated the substitution rate of SARS-CoV-2 from our data (Figure 5). A strong linear correlation was observed between the root-to-tip distance and date of infection, providing strong evidence for a strict molecular clock. Similar to other studies, we estimated an overall substitution rate of 0.00128 substitutions/site/year (95 % confidence interval 0.00126 - 0.00129), or approximately 36 substitutions per year. However, calculated substitution rates for individual VOCs also showed strong linear correlations but ranged from more than 2 times slower for Beta and Omicron to more than ten times slower for Alpha and Delta when compared to the overall substitution rate (Table 4). Clear stepwise accumulations of mutations accompanied each lineage replacement event (Figure 5), leading to stepwise jumps in the number of substitutions fixed within the population over time, and suggesting that selective sweeps are the strongest driver of SARS-CoV-2 evolution.

4. Conclusion

Here, we report a detailed description of the genomic evolution of SARS-CoV-2 on the island community of Reunion based on genomic surveillance data collected between January 2021 and April 2022. While many of the observations we have made are in line with the global evolution of the COVID pandemic, we have seen some interesting aspects that are unique to the insular epidemiology of Reunion Island. Most notably, the dominance of VOC Beta subvariants B.1.351 and B.1.351.2 between weeks 07 and 39 of 2021. We also noted that the interplay between importation rates, epidemic dynamics and substitution rates were all highly contrasted between different VOCs. Overall, the most clear characteristic of SARS-CoV-2 epidemiology is one that is shared globally - that of sequential waves of dominant variants that replace their predecessors.

In common with other island communities, we observed that bottlenecks in strain importation followed by local dissemination resulted in contrasted variant profiles when compared to related continental settings, as well as the amplification and diversification of locally circulating lineages. Within our dataset, the most obvious examples of this came from lineages B.1.622, and Beta variants B.1.351 and B.1.351.2. These observations suggest that insular communities may be important in the emergence of new viral lineages, and it will be important to continue monitoring the genomic evolution of SARS-CoV-2 across the globe.

Acknowledgements

We would like to thank all health professionals and diagnostic laboratories that participated in the sampling and testing of patients, as well as all members of the public health institutions “*Agence Régionale de la Santé*” and “*Santé Publique France*” who helped to coordinate testing and genomic sequencing activities. We would additionally like to thank members from the EMERGEN Consortium (<https://www.santepubliquefrance.fr/emergen>) for their work and support on genomic surveillance in France.

Funding

This work was supported by the «*Délégué Régional Académique à la Recherche et à l'Innovation*» and the “*Agence Régionale de Santé*” through the Genocov project, the European Regional Development Funds through the TFORCE-COVIR program, N°20201437-0027601, and the “Structure fédérative de recherche biosécurité en milieu tropical (BIOST)”.

Authorship contribution statement (CRediT)

DAW: molecular analyses, investigation, methodology, data mining, visualization, writing-original draft; AM: investigation, data mining; MT: molecular analyses; M-A S: molecular analyses; JT: molecular analyses; CL: data analyses, investigation; MC: funding acquisition; M-C J-B: molecular analyses; LJ: data analyses; BL: data analyses; L Y-M: investigation; PM: conceptualization, investigation, funding acquisition, writing-original draft. All authors read and approved the final manuscript.

Conflict of interest

The authors declare no conflict of interest.

Ethics statement

Biological materials were stored and treated using standard diagnostic procedures. No personal data was used in this analysis and samples were anonymized, inhibiting any correlations of samples with the respective patients. Thus, according to national regulations and the institutional rules for Good Scientific Practice, the requirement for submission to an ethical committee and for obtaining patients' informed consent was waived. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

References

1. Emerson BC. Evolution on oceanic islands: molecular phylogenetic approaches to understanding pattern and process. *Mol Ecol.* 2002;11: 951–966. doi:10.1046/J.1365-294X.2002.01507.X

2. Jean K, Burnside WR, Carlson L, Smith K, Guégan JF. An equilibrium theory signature in the island biogeography of human parasites and pathogens. *Glob Ecol Biogeogr.* 2016;25: 107–116. doi:10.1111/GEB.12393
3. Tortosa P, Pascalis H, Guernier V, Cardinale E, Le Corre M, Goodman SM, et al. Deciphering arboviral emergence within insular ecosystems. *Infect Genet Evol.* 2012;12: 1333–1339. doi:10.1016/J.MEEGID.2012.03.024
4. Roth A, Mercier A, Lepers C, Hoy D, Duituturaga S, Benyon E, et al. Concurrent outbreaks of dengue, chikungunya and Zika virus infections - an unprecedented epidemic wave of mosquito-borne viruses in the Pacific 2012-2014. *Euro Surveill.* 2014;19. doi:10.2807/1560-7917.ES2014.19.41.20929
5. Van Bortel W, Dorleans F, Rosine J, Blateau A, Rousseau D, Matheus S, et al. Chikungunya outbreak in the Caribbean region, December 2013 to March 2014, and the significance for Europe. *Euro Surveill.* 2014;19. doi:10.2807/1560-7917.ES2014.19.13.20759
6. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet.* (London, England). 2020;395: 497. doi:10.1016/S0140-6736(20)30183-5
7. Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* (London, England). 2020;395: 565–574. doi:10.1016/S0140-6736(20)30251-8
8. Quick J, Grubaugh ND, Pullan ST, Claro IM, Smith AD, Gangavarapu K, et al. Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nat Protoc.* 2017. doi:10.1038/nprot.2017.066
9. Lambisia AW, Mohammed KS, Makori TO, Ndwiga L, Mburu MW, Morobe JM, et al. Optimization of the SARS-CoV-2 ARTIC Network V4 Primers and Whole Genome Sequencing Protocol. *Front Med.* 2022;9: 836728. doi:10.3389/FMED.2022.836728/FULL
10. O'Toole Á, Scher E, Underwood A, Jackson B, Hill V, McCrone JT, et al. Assignment of epidemiological lineages in an emerging pandemic using the pangolin tool. *Virus Evol.* 2021;7. doi:10.1093/VE/VEAB064
11. Aksamentov I, Roemer C, Hodcroft EB, Neher RA. Nextclade: clade assignment, mutation calling and quality control for viral genomes. *J Open Source Softw.* 2021;6: 3773. doi:10.21105/JOSS.03773
12. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria; 2021. Available: <https://www.r-project.org/>
13. Wickham H, François R, Henry L, Müller K. dplyr: A Grammar of Data Manipulation. 2020. Available: <https://cran.r-project.org/package=dplyr>
14. Paradis E, Schliep K. Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics.* 2019;35: 526–528. doi:10.1093/bioinformatics/bty633

15. Revell LJ. *phytools*: an R package for phylogenetic comparative biology (and other things). *Methods Ecol Evol.* 2012;3: 217–223. doi:10.1111/J.2041-210X.2011.00169.X
16. Yu G, Smith DK, Zhu H, Guan Y, Lam TT. *ggtree*: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. McInerney G, editor. *Methods Ecol Evol.* 2017;8: 28–36. doi:10.1111/2041-210X.12628
17. Wright, Erik S. Using DECIPHER v2.0 to analyze big biological sequence data in R. *R J.* 2016;8: 352. doi:10.32614/RJ-2016-025
18. Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol Biol Evol.* 2015;32: 268. doi:10.1093/MOLBEV/MSU300
19. Sagulenko P, Puller V, Neher RA. TreeTime: Maximum-likelihood phylodynamic analysis. *Virus Evol.* 2018;4. doi:10.1093/VE/VEX042
20. Shu Y, McCauley J. GISAID: Global initiative on sharing all influenza data – from vision to reality. *Eurosurveillance.* 2017. doi:10.2807/1560-7917.ES.2017.22.13.30494
21. Mercier A, Wilkinson DA, Lebarbenchon C, Mavingui P, Yemadje-Menudier L. Spread of SARS-CoV-2 Variants on Réunion Island, France, 2021. *Emerg Infect Dis.* 2022;28: 895. doi:10.3201/EID2804.212243
22. Hodcroft EB, Zuber M, Nadeau S, Vaughan JC, Crawford KHD, Althaus CL, et al. Spread of a SARS-CoV-2 variant through Europe in the summer of 2020. *Nat* 2021 5957869. 2021;595: 707–712. doi:10.1038/s41586-021-03677-y
23. Plante JA, Liu Y, Liu J, Xia H, Johnson BA, Lokugamage KG, et al. Spike mutation D614G alters SARS-CoV-2 fitness. *Nat* 2020 5927852. 2020;592: 116–121. doi:10.1038/s41586-020-2895-3
24. Nelde A, Bilich T, Heitmann JS, Maringer Y, Salih HR, Roerden M, et al. SARS-CoV-2-derived peptides define heterologous and COVID-19-induced T cell recognition. *Nat Immunol* 2020 221. 2020 22: 74–85. doi:10.1038/s41590-020-00808-x
25. Chu DKW, Hui KPY, Gu H, Ko RLW, Krishnan P, Ng DYM, et al. Introduction of ORF3a-Q57H SARS-CoV-2 Variant Causing Fourth Epidemic Wave of COVID-19, Hong Kong, China. *Emerg Infect Dis.* 2021;27: 1492. doi:10.3201/EID2705.210015
26. Azad GK, Khan PK. Variations in Orf3a protein of SARS-CoV-2 alter its structure and function. *Biochem Biophys Reports.* 2021;26: 100933. doi:10.1016/J.BBREP.2021.100933
27. Harvey WT, Carabelli AM, Jackson B, Gupta RK, Thomson EC, Harrison EM, et al. SARS-CoV-2 variants, spike mutations and immune escape. *Nat Rev Microbiol* 2021 197. 2021;19: 409–424. doi:10.1038/s41579-021-00573-0
28. Davis JJ, Long SW, Christensen PA, Olsen RJ, Olson R, Shukla M, et al. Analysis of the ARTIC Version 3 and Version 4 SARS-CoV-2 Primers and Their Impact on the Detection of the G142D Amino Acid Substitution in the Spike Protein. *Microbiol Spectr.* 2021;9. doi:10.1128/SPECTRUM.01803-21/ASSET/859C34AA-540F-453B-9BAA-

36FA4E168866/ASSETS/IMAGES/MEDIUM/SPECTRUM.01803-21-F001.GIF

29. Ciuffreda L, González-Montelongo R, Alcoba-Florez J, García-Martínez de Artola D, Gil-Campesino H, Rodríguez-Pérez H, et al. Tracing the trajectories of SARS-CoV-2 variants of concern between December 2020 and September 2021 in the Canary Islands (Spain). *Front Cell Infect Microbiol.* 2022;12: 919346. doi:10.3389/fcimb.2022.919346
30. Viala VL, Slavov SN, de Lima LPO, Lima ARJ, Ribeiro G, Martins AJ, et al. The Divergent Pattern of SARS-CoV-2 Variant Predominance and Transmission Dynamics in the Brazilian Island of Ilhabela. *Viruses.* 2022;14. doi:10.3390/v14071481
31. Althaus CL, Baggio S, Reichmuth ML, Hodcroft EB, Riou J, Neher RA, et al. A tale of two variants: Spread of SARS-CoV-2 variants Alpha in Geneva, Switzerland, and Beta in South Africa. *medRxiv.* 2021; 2021.06.10.21258468. doi:10.1101/2021.06.10.21258468
32. Hart WS, Miller E, Andrews NJ, Waight P, Maini PK, Funk S, et al. Generation time of the alpha and delta SARS-CoV-2 variants: an epidemiological analysis. *Lancet Infect Dis.* 2022;22: 603–610. doi:10.1016/S1473-3099(22)00001-9/ATTACHMENT/859BF3E7-B49C-4904-827D-D6878A8E429A/MMC2.XLSX

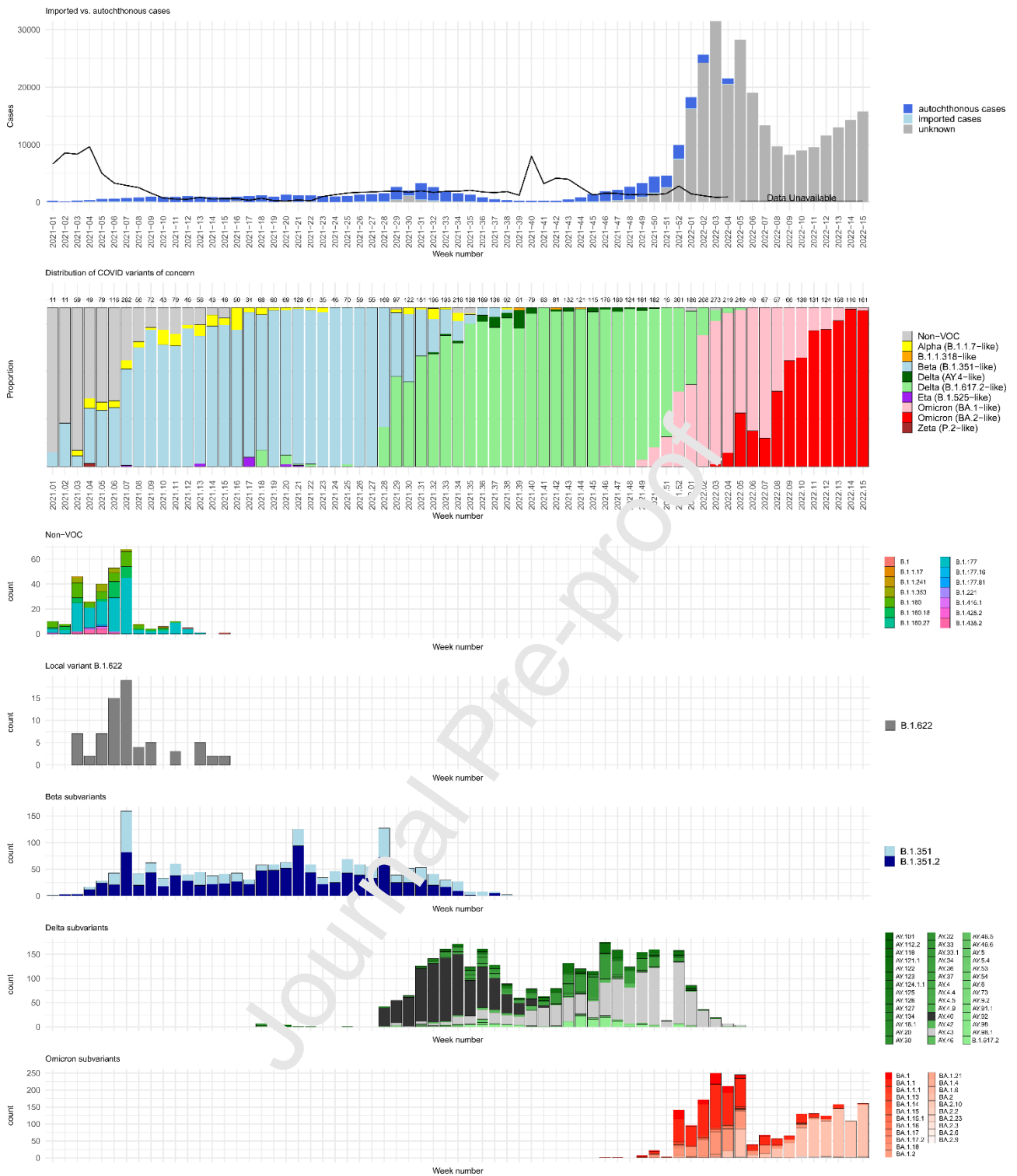


Figure 1: Summary of the epidemiological and genomic classification data from Reunion 2021-2022. Panels display (from top to bottom): 1) Weekly numbers of reported cases, and the ratio between imported and autochthonous cases (black line), 2) Weekly proportions of genomes attributed to different variants of concern, 3-7) Weekly numbers of sequenced genomes attributed to different PANGO lineages, data are displayed separately for each major variant of concern.

| #Mutations (Median) [IQR] | Wuhan | B.1.474 | B.1.622 | B.1.629 |
|---------------------------------|-------|---------------|---------------|-------------|
| Wuhan | 0 | | | |
| B.1.474 | 25 | 8 [5,11] | | |
| B.1.622 | 24 | 33 [31,36] | 1 [1,2] | |
| B.1.629 | 33 | 38 [35,41] | 33 [32,35] | 9 [7,11] |

Table 1: Numbers of mutations between lineages related to B.1.622.

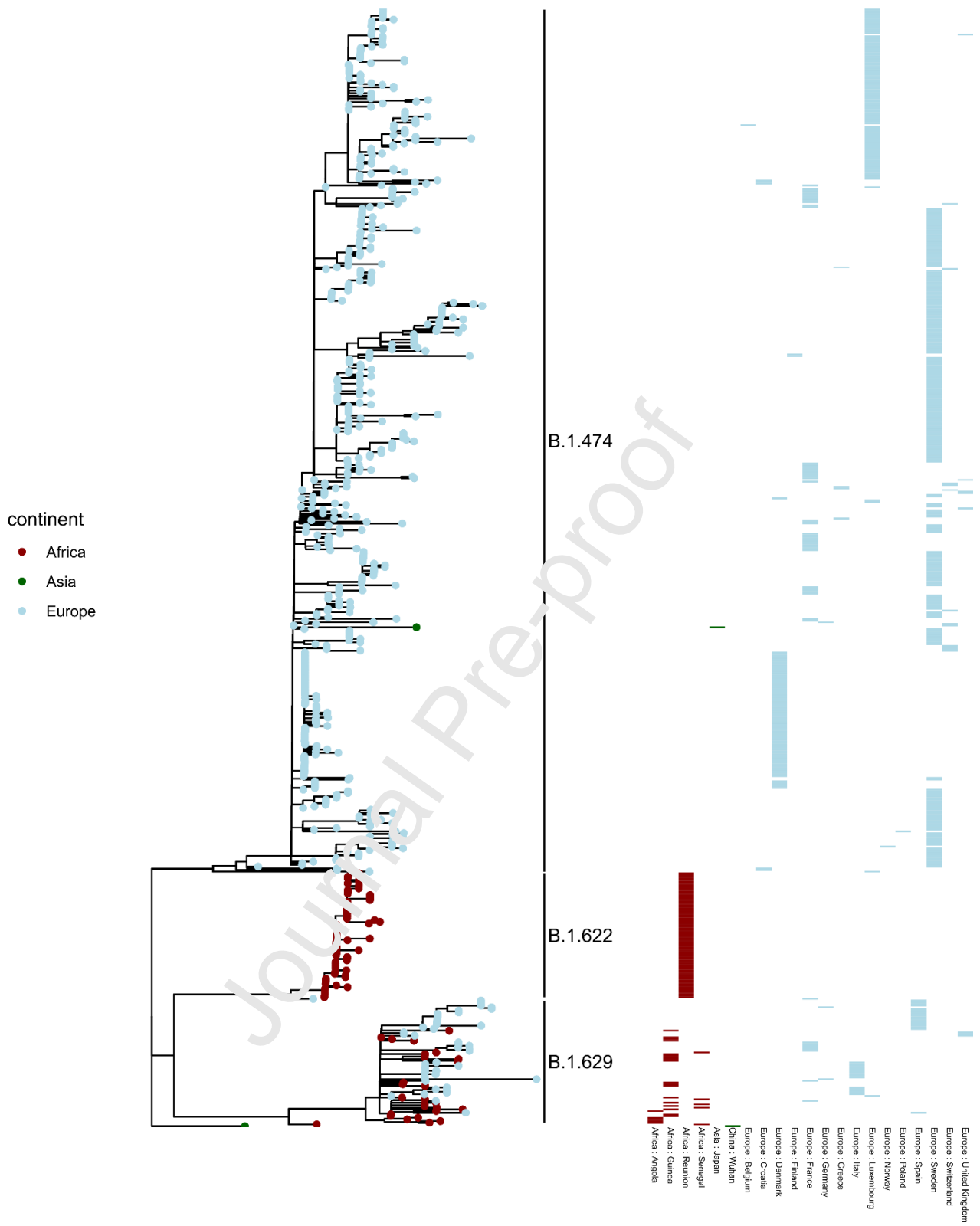


Figure 2: Comparison of lineage B.1.622 to its closest genetic relatives in GISAID.

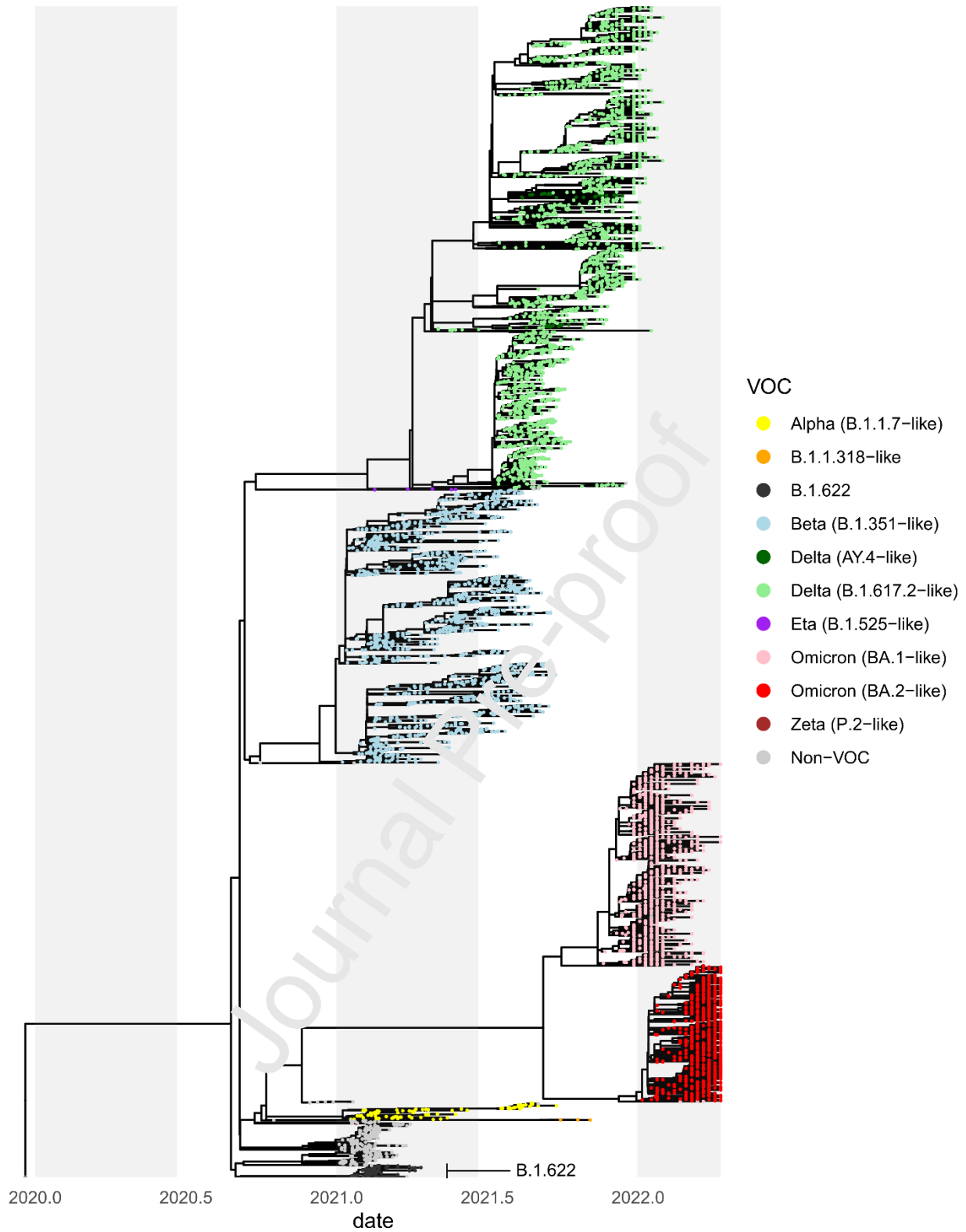


Figure 3: Phylogenetic analyses of SAR-CoV-2 genomes from Reunion 2021-2022. a) Timetree phylogeny, leaves are coloured by variant of concern, shaded regions represent 6-month intervals.

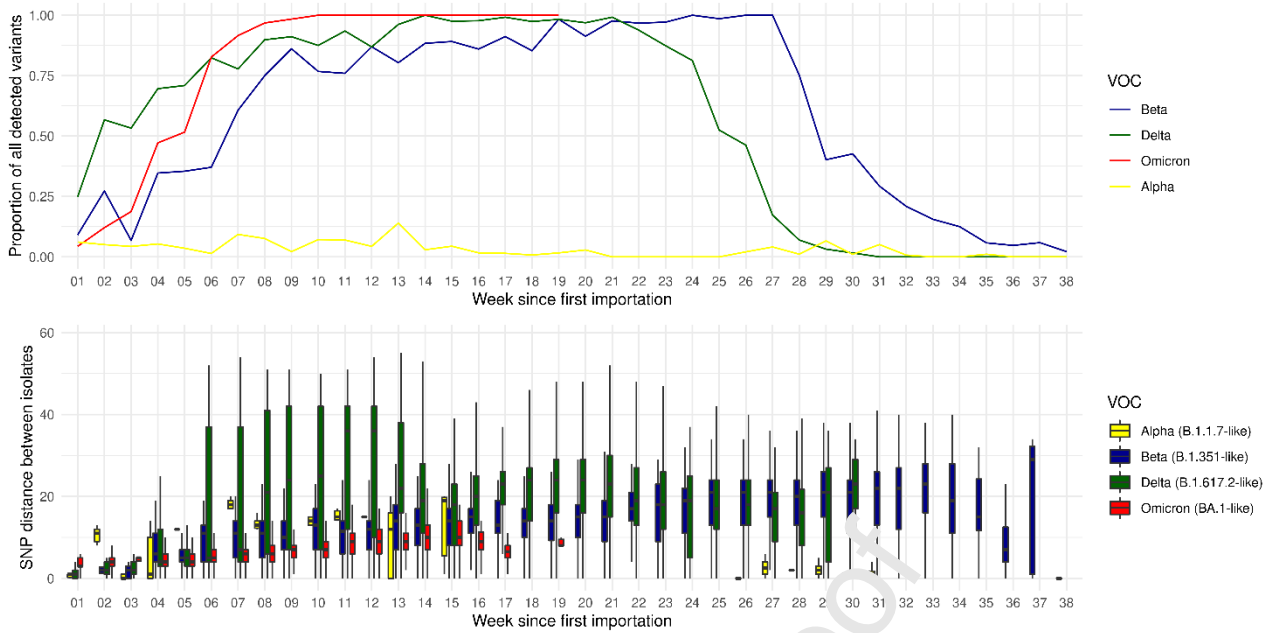


Figure 4: Dynamics of VOCs in Reunion 2021-2022. TOP) Proportion of all sequenced isolates belonging to each VOC. BOTTOM) Total measured genetic diversity of main lineages over time. Boxplots show the distribution in the number of mutations between isolates belonging to the same VOC sequenced on the same week. Throughout, week number is measured relative to the start of each “wave” of variant circulation, as follows: Beta started in week one of 2021, Delta started in week 27 of 2021 and Omicron started in week 42 of 2021, the first observation of Alpha was in week 3 of 2021.

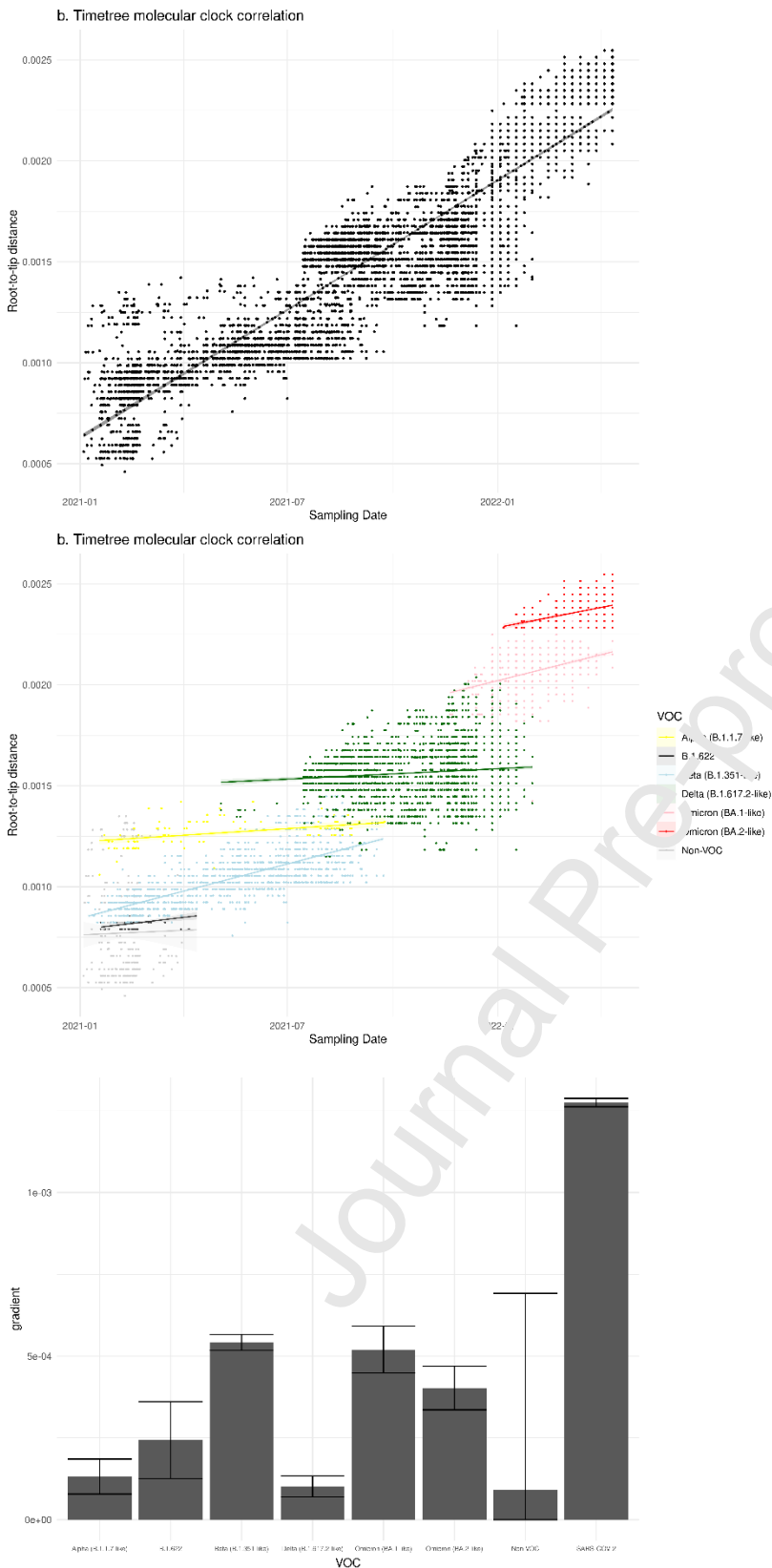


Figure 5: Substitution rate estimates for SARS-CoV-2 genomes from Reunion 2021-2022. a) Overall linear correlation between root-to-tip distance in phylogenetic analyses and collection date for all SARS-CoV-2 genomes. b) Individual correlations between root-to-tip distance and collection date for all variants of concern. c) Average substitutions per site per year calculated for each major variant of concern, error bars represent the 95% confidence intervals of the linear regression for each correlation.

| Substitution | BETA | | | DELTA | | OMICRON | | |
|--------------|---------|---------|-----------|-------|-------|---------|--------|------|
| | B.1.622 | B.1.351 | B.1.351.2 | AY.40 | AY.43 | BA.1 | BA.1.1 | BA.2 |
| N:Q70R | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| N:S187L | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF1a:H110Y | 0.9 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF1a:L3977V | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF1a:M3621I | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF1a:S428N | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF1a:T2196N | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF1b:S2312L | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF3a:Q57H | 1.0 | 1.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF3a:S216P | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF3a:T14I | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF9b:K67E | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| E:P71L | 0.0 | 1.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| N:T205I | 0.0 | 1.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF1a:K1655N | 0.0 | 1.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF1a:K3353R | 0.0 | 1.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF1a:T265I | 0.0 | 1.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF3a:S171L | 0.0 | 1.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| S:D215G | 0.0 | 1.0 | 0.9 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| S:D80A | 0.0 | 1.0 | 0.9 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| S:K417N | 0.0 | 1.0 | 0.9 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| S:E484K | 0.0 | 0.9 | 0.8 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| S:A701V | 0.0 | 1.0 | 0.9 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| M:I82T | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| N:D377Y | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| N:D63G | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| N:R203M | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| ORF1b:G662S | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| ORF1b:P1000L | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 | 1.0 | 0.0 |
| ORF3a:S26L | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| ORF7a:T120I | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF7a:V82A | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| ORF9b:T60A | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| S:D950N | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| S:L452R | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| S:P681R | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| S:R158G | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| S:T19R | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| S:T478K | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| E:T9I | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| M:A63T | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| M:Q19E | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| N:G204R | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| N:P13L | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| N:R203K | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| ORF1a:P3395H | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| ORF1a:T3255I | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| ORF1b:I1566V | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| ORF9b:P10S | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| S:D796Y | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| S:E484A | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| S:H655Y | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| S:N501Y | 0.0 | 0.9 | 0.8 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| S:N679K | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| S:N969K | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| S:P681H | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| S:Q493R | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| S:Q498R | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| S:Q954H | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| S:G339D | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| S:S373P | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| S:S375F | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| S:S477N | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| S:Y505H | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| S:N764K | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.7 | 0.8 | 1.0 |

Table 2: Amino acid substitutions conserved within VOCs that circulated in Reunion 2021-2022

| | BETA | DELTA | OMICRON |
|--|------|-------|---------|
|--|------|-------|---------|

| | B.1.622 | B.1.351 | B.1.351.2 | AY.40 | AY.43 | BA.1 | BA.1.1 | BA.2 |
|--------------|---------|---------|-----------|-------|-------|------|--------|------|
| ORF1a:A3456V | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF1a:N2596S | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| S:L18F | 0.0 | 0.1 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| N:L139F | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| N:R385K | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF1a:A3209V | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF1a:H2092Y | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF1a:I431M | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF1a:P1640L | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF1a:P309L | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF1a:V3718A | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF1b:H2285Y | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF1b:M809V | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF7a:L116F | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF8:S69L | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ORF9b:I5T | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| S:D574Y | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| N:G215C | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| N:Q9L | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| ORF1a:A1306S | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| ORF1a:P2046L | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| ORF1a:P2287S | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| ORF1a:T3646A | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| ORF1a:V2930L | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| ORF1b:A1918V | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| ORF1b:L829I | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| ORF7b:T40I | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| ORF9b:S6C | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| S:G142D | 0.0 | 0.0 | 0.0 | 0.0 | 0.9 | 0.0 | 0.0 | 1.0 |
| M:D3G | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 |
| ORF1a:A2710T | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 |
| ORF1a:I3758V | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 |
| ORF1a:K856R | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 |
| ORF1a:L2084I | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 |
| S:A67V | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 |
| S:G496S | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 |
| S:L212I | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 |
| S:L981F | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 |
| S:N856K | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 |
| S:S371L | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 |
| S:T547K | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 |
| S:T95I | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 |
| S:Y145D | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 |
| N:S413R | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| ORF1a:G1307S | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| ORF1a:L3027F | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| ORF1a:L3201F | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| ORF1a:S135R | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| ORF1a:T3090I | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| ORF1a:T842I | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| ORF1b:R1315C | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| ORF1b:T2163I | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| ORF3a:T223I | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| ORF6:D61L | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| S:A27S | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| S:D405N | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| S:N440K | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| S:R408S | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| S:S371F | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| S:T19I | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| S:T376A | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| S:V213G | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| S:R346K | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 |

Table 3: Amino acid substitutions conserved within PANGO lineages that circulated in Reunion 2021-2022

| Lineage | Substitutions per site per year | R ² | 95 % CI | Relative substitution rate compared to All SARS-CoV-2 |
|---------|---------------------------------|----------------|---------|-------------------------------------------------------|
|---------|---------------------------------|----------------|---------|-------------------------------------------------------|

| | | | | |
|------------------------|-----------|-------|------------------------|-------|
| Alpha (B.1.1.7-like) | 0.000132 | 0.435 | [0.0000781 - 0.000185] | 0.103 |
| B.1.622 | 0.000243 | 0.444 | [0.000125 - 0.000361] | 0.190 |
| Beta (B.1.351-like) | 0.000542 | 0.736 | [0.000518 - 0.000566] | 0.423 |
| Delta (B.1.617.2-like) | 0.000102 | 0.114 | [0.0000698 - 0.000134] | 0.080 |
| Omicron (BA.1-like) | 0.000520 | 0.372 | [0.000449 - 0.000591] | 0.406 |
| Omicron (BA.2-like) | 0.000402 | 0.378 | [0.000336 - 0.000469] | 0.314 |
| Non-VOC | 0.0000913 | 0.018 | [0 - 0.000692] | 0.071 |
| All SARS-CoV-2 | 0.00128 | 0.915 | [0.00126 - 0.00129] | 1 |

Table 4. Substitution rates for major SARS-CoV-2 lineages in Reunion 2021-2022.

Conflict of interest

The authors declare no conflict of interest.

Journal Pre-proof

Credit author statement :

David A Wilkinson : Conceptualization, Methodology, Software, Validation, Formal analysis, Data Curation, Writing – Original Draft, Writing – Review and Editing, Visualization, Supervision, Project administration

Alizé Mercier : Methodology, Investigation, Data Curation, Writing – Original Draft, Writing – Review and Editing

Magali Turpin : Formal analysis, Investigation

Marie-Alice Simbi : Formal analysis, Investigation

Jonathan Turpin : Formal analysis, Investigation

Camille Lebarbenchon : Formal analysis, Investigation, Writing – Original Draft

Maya Cesari : Supervision, Project administration, Funding acquisition

Marie-Christine Jaffar-Bandjee : Investigation, Resources

Laurence Josset : Investigation, Writing – Original Draft

Luce Yemadie-Menudier : Resources, Investigation, Project administration

Bruno Lina: Resources, Investigation, Project administration, Writing – Original Draft

Patrick Mavingui: Conceptualization, Formal analysis, Writing – Original Draft, Resources, Supervision, Project administration, Funding acquisition

Highlights

- The epidemiology and evolution of COVID-19 in Reunion Island was unique due to its insular context.
- Globally uncommon Beta sublineages were majority variants on the island during much of 2021.
- Substitution rates were heterogeneous between SARS-CoV-2 lineages.
- Near-complete selective sweeps were the main drivers of COVID epidemiology and evolution.

Journal Pre-proof